Isolated Nucleotide Sequences Responsible for the Tomato *High Pigment-1* Mutant Phenotypes (*hp-1 and hp-1^{w}*) and Uses thereof

Field of the Invention

The present invention relates to modified nucleotide sequences that are responsible for producing the *high pigment-1* and *high pigment-1^{w}* phenotypes in tomatoes. More specifically, the present invention discloses point mutations within the tomato homolog of the *Arabidopsis thaliana* and human *DDB1* (UV damaged DNA binding protein 1) gene, and the uses of said modified nucleotide sequences.

Background of the Invention

Plants respond to light intensity, direction, duration, and spectral quality by modulating their developmental processes in an array of interactions that are referred to as photomorphogenesis. Photomorphogenic mutants have been proven to be an excellent tool in research of the complex interactions between light and plant development and some of them have also been used in several agricultural crop breeding programs. Photomorphogenic mutants have been reported in a number of species, including *Arabidopsis*, *Sorghum*, *Brassica*, tobacco, tomato and pea. In general, these mutants may be classified either as defective in photoreceptors, or altered in some aspect of light signal transduction chain (Chory, 1993).

Several photomorphogenic mutants have been described in tomato (*lycopersicon esculentum*). Among these, mutants carrying the monogenic recessive *high pigment* (*hp-1*, *hp-1^{w}*, *hp-2*, and *hp-2^{j}*) and *dark green* (*dg*) mutations are characterized by their exaggerated light responsiveness.

These mutants display higher anthocyanin levels, shorter hypocotyls, and greater fruit pigmentation in comparison to their semi-isogenic wild type plants (Mochizuki and Kamimura 1984; Wann et al. 1985). The increased fruit pigmentation seen in these mutants is due to significantly elevated levels of carotenoids, primarily lycopene, and flavonoids in the mature ripe red fruit. As a consequence of their effect on fruit color, *hp* and *dg* mutations were introgressed into several commercial processing and fresh-market tomato cultivars that are currently marketed as Lycopene Rich Tomatoes (LRT) (Wann, 1997).

The *hp-1* mutant was originally discovered as a spontaneous mutant in 1917 at the Campbell Soup Company farms (Riverton, NJ) (Reynard, 1956), The $hp\text{-}1^{w}$ mutant appeared among progeny of a plant raised from ethyl methanesulfonate (EMS)-treated seeds of the genotype GT (Peters et al. 1989), the *hp-2* mutant was reported in the Italian San Marzano variety in 1975 (Soressi 1975), the $hp\text{-}2^{j}$ mutant was found among progeny of a T-DNA-transformed plant (cv Moneymaker) (van Tuinen et al. 1997), and the *dg* mutant appeared in trellised planting of the Manapal variety (Konsler 1973). Despite some initial confusion, it is now clear that there are two *HP* genes - *HP-1* and *HP-2* - in the tomato genome, that map to chromosomes 2 and 1, respectively (van Tuinen et al. 1997; Yen et al. 1997). (Van Tuinen et al. 1997; Yen et al. 1997). At each of these loci, two of the above mentioned mutant alleles have been initially identified: *hp-1* and $hp\text{-}1^{w}$, *hp-2* and $hp\text{-}2^{j}$ (Kerckhoff and Kendrick 1997; Van Tuinen et al. 1997).

WO 99/29866 discloses the cloning and sequencing of the *HP-2* gene, said gene being found to encode the tomato homolog of the *Arabidopsis* nuclear protein *DEETIOLATED1* (*DET1*).

This publication further discloses that a point mutation and deletion mutation, both of which are located in exon 11 at the 3′ end of the coding sequence of *HP-2*, give rise to the previously-identified *hp-2^{j}* and *hp-2* mutants respectively. In the case of the *hp-2* mutant, a point mutation directs alternative splicing of intron 10 that leads to a nine base pair deletion in exon 11.

Co-owned WO 03/57917 discloses another point mutation in the tomato homolog of the *Arabidopsis DET1* gene that is responsible for the *dg* mutation, and which therefore comprises a 3rd mutant allele at the *HP-2* locus.

It is a purpose of the present invention to provide isolated nucleotide sequences containing the mutations responsible for the *high pigment-1* (*hp-1*) and *high pigment-1^{w}* (*hp-1^{w}*) photomorphogenic mutants of tomato plants.

It is a further purpose of the present invention to provide DNA markers that may be used as a molecular diagnostic tool for the identification and selection of *hp-1* and *hp-1^{w}* mutants.

A yet further purpose of the present invention is to provide molecular diagnostic tools that may be used for genotypic selection in the production of lycopene-enhancing double mutants.

Other purposes and advantages of the present invention will become apparent as the description proceeds.

Summary of the Invention

It has now been found that the mutations responsible for both the *hp-1* and *hp-1*^{w} mutant phenotypes are located within the tomato homolog of the human and *Arabidopsis thaliana* UV DAMAGED DNA BINDING Protein 1 (*DDB1*) gene.

The present invention is primarily directed to isolated nucleotide sequences responsible for the tomato *hp-1* and *hp-1*^{w} phenotypes, wherein each of said sequences comprises an altered tomato *DDB1* gene sequence or fragment or homolog thereof. In the case of the *hp-1* mutation, the alteration in said sequence or fragment or homolog comprises a single A^{931}-to-T^{931} base transversion in the tomato homolog of the *DDB1* coding sequence. In the case of the *hp-1*^{w} mutation, the alteration in said sequence or fragment or homolog comprises a single G^{2392}-to-A^{2392} transition in the tomato homolog of the *DDB1* coding sequence.

In one preferred embodiment of the present invention, the isolated nucleotide sequence encoding the *hp-1* mutation comprises the sequence defined as SEQ ID NO:1 in the sequence listing. It is to be noted that all of the sequences contained in the enclosed sequence listing are to be considered to form an integral part of the present disclosure.

In another preferred embodiment of the present invention, the isolated nucleotide sequence encoding the *hp-1*^{w} mutation comprises the sequence defined as SEQ ID NO:2 in the sequence listing.

It is to be understood that the present invention also includes within its scope all fragments of the above-defined sequences that encode the *hp-1* and hp-1^w mutations, wherein said fragments comprise the region of the *DDB1* gene sequence containing the mutated nucleotide, that is, the region containing nucleotide 931 in the case of the *hp-1* mutation, and nucleotide 2392 in the case of the *hp-*1^w mutation.

The present invention is also directed to methods for detecting the presence of the *hp-1* and (independently) the *hp-*1^w mutations in plant material.

Thus, in one embodiment of this aspect, the present invention provides a method for detecting the presence of the *hp-1* mutation in a plant, comprising the steps of isolating the genomic DNA from said plant, amplifying a gene fragment containing said *hp-1* mutation from said genomic DNA by use of a PCR technique and determining the presence of said *hp-1* mutation in said genomic DNA.

In another embodiment, the present invention provides a method for detecting the presence of the *hp-*1^w mutation in a plant, comprising the steps of isolating the genomic DNA from said plant, amplifying a gene fragment containing said *hp-*1^w mutation from said genomic DNA by use of a PCR technique and determining the presence of said *hp-*1^w mutation in said genomic DNA

Any suitable technique may be used to determine the presence of the *hp-1* and (independently) the *hp-*1^w mutations in the plant material. However, in a preferred embodiment, the presence of said mutations is determined by the use of a pyrosequencing technique, wherein the sequence data obtained

from said technique is compared with the sequences defined in SEQ ID NO:1 (in the case of *hp-1*) and SEQ ID NO:2 (in the case of *hp-1*^{w}).

In a particularly preferred embodiment, the above-defined method of determining the presence of the *hp-1* and (independently) the *hp-1*^{w} mutations is applied to material obtained from the species *Lycopersicon esculentum*.

In one particularly preferred embodiment, the above-disclosed method of determining the presence of the *hp-1* and (independently) the *hp-1*^{w} mutations is used as a means of quality control, or post-control in seed production, for detecting the presence of the *dg* allele in cultivars and their parental lines. The term post-control is used herein to indicate quality control checks that are performed following seed production, in order to confirm the intended genotype of said seeds.

In another aspect, the present invention is also directed to a method for the determination of the presence of two different photomorphogenic mutations in a plant, wherein one of said mutations is either the *hp-1* or the *hp-1*^{w} mutation, comprising detecting the presence of a photomorphogenic mutation other than the *hp-1* or the *hp-1*^{w} mutation by either genotypic or phenotypic selection means, and detecting the presence of the *hp-1* or the *hp-1*^{w} mutation by means of the method disclosed hereinabove. In one preferred embodiment of this aspect of the invention, the phenotypic selection means for determining the presence of the non-*hp-1*, non-*hp-1*^{w} photomorphogenic mutation comprises germinating seeds obtained from the plant in which the presence of the mutations is being determined in a temperature controlled chamber, under a yellow plastic screen that is opaque to

light having a wavelength less than 500nm, and selecting non-etiolated seedlings.

The present invention is also directed to a method for preparing double-mutant lines of *Lycopersicon esculentum* having genotype *hp-1/hp-1 p/p*, wherein *p* represents any recessive photomorphogenic lycopene-enhancing mutation that is genetically unlinked to the *hp-1* mutation, said method comprising the steps of:

a) cross-hybridization of a homozygous *hp-1/hp-1* line or plant with a homozygous *p/p* line or plant to yield double heterozygous *hp-1/+ p/+* F_1 plants;
b) self-crossing of the F_1 plants obtained in step (a) in order to yield F_2 seeds;
c) identification of double homozygous plants *hp-1/hp-1 p/p* by means of the application of the method defined in claim 7 and a method for detecting the presence of the *p* mutation;
d) self-crossing of the double homozygous plants identified in step (c) to generate F_3 seeds, and germination of said seeds.

In one preferred embodiment of this aspect of the invention, the mutation *p* is the *dg* mutation. In this case, the determination of the presence of the *dg* mutation in step (c) of the method may be performed using the marker for the *dg* mutation disclosed in co-owned, co-pending application PCT/IL03/00023.

Similarly, the present invention is also directed to a method for preparing double-mutant lines of *Lycopersicon esculentum* having genotype *hp-1^w/hp-1^w p/p*, wherein *p* represents any recessive photomorphogenic lycopene-enhancing mutation that

is genetically unlinked to the *hp-1^w* mutation, said method comprising the steps of:

a) cross-hybridization of a homozygous *hp-1^w/hp-1^w* line or plant with a homozygous *p/p* line or plant to yield double heterozygous *hp-1^w/+ p/+* F_1 plants;
b) self-crossing of the F_1 plants obtained in step (a) in order to yield F_2 seeds;
c) identification of double homozygous plants *hp-1^w/hp-1^w p/p* by means of the application of the method defined in claim 10 and a method for detecting the presence of the *p* mutation;
d) self-crossing of the double homozygous plants identified in step (c) to generate F_3 seeds, and germination of said seeds.

In one preferred embodiment of this aspect of the invention, the mutation *p* is the *dg* mutation. In this case, the determination of the presence of the *dg* mutation in step (c) of the method may be performed using the marker for the *dg* mutation disclosed in co-owned, co-pending application PCT/IL03/00023.

The present invention also encompasses within its scope double-mutant hybrid plants of the species *Lycopersicon esculentum* having genotype *hp-1/hp-1 p/p and (independently) hp-1^w/hp-1^w p/p* prepared by the above-disclosed methods.

All the above and other characteristics and advantages of the present invention will be further understood from the following illustrative and non-limitative examples of preferred embodiments thereof.

Brief Description of the Drawings

Fig. 1. Nucleotide sequence of the genomic fragment used to design pyrosequencing primers for the *hp-1* mutation (the single nucleotide polymorphism is in underlined large bold letters, the forward and the reverse primers are underlined and the sequencing primer is in italic).

Fig. 2. Partial mapping results of the tomato *DDB1* gene (map of the tomato chromosome 2, showing the location of the *HP-1* gene (*hp*) was adopted from Yen et al. (1997)).

Fig. 3. Typical pyrosequencing genotyping results for the *hp-1* mutation at the *DDB1* locus (Because of the reverse orientation of the sequencing primer, the mutant genotype is characterized by A and the normal genotype by T).

Fig. 4. Partial ClustalW protein Alignment of DDB1 showing the location of the *hp-1* (a) and *hp-1*^{w} (b) amino-acid substitutions [presented are *Arabidopsis* DDB1A (At_DDB1A=NP_192451), *Arabidopsis* DDB1B (At_DDB1B=NP_193842), tomato cv. Ailsa Craig (Le= AY452480), rice (Os=BAB20761), human (Hs=DDB1_Human), *Drosophila* (Dm=XP_081186), chicken (Gg=BAC56999), and *S. pombe* (Sp= NP_593580)]. Identical residues are black shaded whereas similar residues are gray shaded.

Fig. 5. Complete nucleotide coding sequence of the normal wild-type tomato *DDB1* gene (the start, ATG, and the termination, TAG, codons are underlined. Location of A^{931} and G^{2392}, whose transversion and transition leads to the *hp-1* and *hp-1*^{w} phenotypes, respectively, are in large bold letters).

Fig. 6. Complete amino-acids sequence of the normal wild-type tomato *DDB1* gene (Asparagine^{311} and Glutamic-acid^{798} whose substitution leads to the *hp-1* and *hp-1^{w}* phenotypes, respectively, are in large bold letters).

Detailed Description of Preferred Embodiments

In one of its aspects, as described hereinabove, the present invention provides a method for detecting the presence of the *hp-1* and *hp-1^{w}* mutations in a plant.

In the case of the *hp-1* mutation, this method comprises the steps of isolating a genomic DNA fragment comprising the region of the *DDB1* gene containing the site of the single nucleotide polymorphism (SNP) responsible for the *hp-1* phenotype (at nucleotide position 931), cloning said fragment, sequencing said cloned fragment and determining the presence of the *hp-1* mutation by means of detecting the A/T transversion at position 931 of the sequenced genomic fragment.

In a particularly preferred embodiment of this method of the present invention, the sequencing of the cloned fragment is achieved by means of a pyrosequencing reaction. Prior to the pyrosequencing reaction, the SNP-containing genomic fragment is amplified by means of a PCR technique, the details of which will be described hereinbelow.

The term "PCR" (or polymerase chain reaction) technique as used hereinabove and hereinbelow refers to a family of techniques that are based on the use of heat-stable polymerases for achieving the amplification (i.e. increase in number of copies) of specific DNA sequences by repeated polymerase reactions. This reaction can be used as a replacement for cloning: all that is required is knowledge

of the nucleic acid sequence. In order to carry out PCR, primers are designed which are complementary to the sequence of interest. The primers are then generated by automated DNA synthesis.

PCR and other methods of amplifying DNA and/or RNA are well known in the art, and can be used according to the present invention without the need for undue experimentation, based on the teaching and guidance presented herein. Several PCR methods (as well as related techniques) are described, for example, in US patent Nos. 4,683,195, 4,683,202, 4,800,159, 4,965,188, as well as in Innis *et al.* eds., *PCR Protocols: A guide to method and applications*.

The following examples are provided for illustrative purposes and in order to more particularly explain and describe the present invention. The present invention, however, is not limited to the particular embodiments disclosed in these examples.

EXAMPLES

Materials and methods

Plant material and crosses

Seeds from the normal, open pollinated, tomato *(Lycopersicon esculentum)* cv. Ailsa Craig and a line nearly isogenic and homozygous for the *hp-1* mutation were kindly provided by J.J. Giovannoni, of the Boyce Thompson Institute for Plant Research, Ithaca, NY, USA.

Seeds from cv. Rutgers homozygous for the a *hp-1* mutation (LA3004), as well as seeds from $hp\text{-}1^{w}/hp\text{-}1^{w}$ mutant plants and their isogenic normal plants in GT background (LA LA4012 and LA4011, respectively), were provided by R.T.

Chetelat, of the Tomato Genetics Cooperative, UC Davis, CA, USA. The genotype GT is a tomato breeding line, resistant to mosaic virus, and similar in morphology to cv. Moneymaker, originally obtained from Deruiterzonen, Bleiswijk, the Netherlands (Koornneef et al. 1990). The *hp-1^{w}/hp-1^{w}* mutant plants appeared among progeny of a plant raised from EMS-treated seeds of the genotype GT (Peters et al. 1989). Therefore, these plants are highly isogenic to the normal GT genotype. Mutant *hp-1^{w}/hp-1^{w}* plants show a more extreme phenotype compared to *hp-1/hp-1* plants, and it was clearly shown that *hp-1* and *hp-1^{w}* are allelic (Peters et al. 1989).

A processing *hp-1/hp-1* mutant hybrid, LRT89, two *hp-1/hp-1* breeding lines, L525 and L527, and a normal breeding line, N671, were developed by the late R. Frankel, D. Lapushner and I. Levin at the Volcani Center. Seeds from two *hp-1/hp-1* processing hybrids, HA3501 and HA3502, developed by Hazera Genetics Inc., Israel, were provided by Mr. Ezri Peleg. Seeds of the heterozygous *hp-1/+* cultivar, cv. 124, were also provided by Hazera Genetics Inc., Israel. Several normal +/+ tomato cultivars used in this study, i.e.: Moneymaker, M82, Brigade, VF-36, 189, Manapal, NC8288 and Florida, were from seed stocks available at the Volcani center. DNA was also extracted from single plants of AB427, AB510 and AB747, three *hp-1/hp-1* processing hybrids developed by AB Seeds Inc., Israel.

Normal cv. Ailsa Craig plants were crossed with their nearly isogenic *hp-1* mutant plants to yield F_1 seeds. These F_1 plants were allowed to self pollinate to yield F_2 seeds. A sample of 123 F_2 seedlings was used for the linkage analysis carried out in this study.

Genomic DNA extraction and Southern blot hybridization

Genomic DNA was extracted from individual plants according to Fulton et al. 1995. To determine the copy number of the *DDB1* gene in the tomato genome, Southern blot hybridization was carried out according to the following procedure: Genomic DNA extracted from both *L. esculentum* (cv. M82) and *L. pennellii* was digested with *Eco*R I, *Eco*R V, *Dra* I, *Hae* III, *Sca* I, and *Mva* I restriction endonucleases. Following electrophoresis in 1.0% agarose gel and Southern transfer, the DNA was hybridized with a P^{32} labeled DNA probe containing 1346 bp of the 5' coding sequence of the *DDB1* gene. Southern blot transfer and DNA hybridization were done according to Levin and Smith (1990).

Design of PCR primers

Sequence analysis and locus-specific primer design were carried out with the DNAMAN, Sequence Analysis Software version 4.1 (Lynnon BioSoft, Quebec, Canada). All DNA primers used were purchased from M.B.C Molecular Biology Center Ltd., Ness-Ziona, Israel.

PCR reactions

PCR reactions were used for mapping, cloning and amplification of DNA products for direct sequencing and pyrosequencing. For all of these purposes, the amplification reactions (25 ml final volume) were performed with 10 ng template DNA, 25 mM TAPS (pH=9.3 at 25°C), 50 mM KCl, 2mM MgCl2, 1mM B-mercaptoethanol, 0.2 mM of each of the four deoxyribonucleotide triphosphates (dATP, dCTP, dGTP and dTTP), 10 pmoles of each of two primers and 1 unit of thermostable *Taq* DNA polymerase (SuperNova *Taq* polymerase, Madi Ltd., Rishon Le Zion, Israel). Reactions were carried out in an automated thermocycler (MJ Research Inc., Watertown, MA, USA).

For mapping and direct sequencing, initial incubation was at 94°C for 3 min, followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 58°C for 30 sec, and polymerization at 72°C for 1-2 min, depending upon the PCR product size. Final polymerization at 72°C was carried out, for 5 min, after completion of the above cycles. The PCR amplification products were visualized by electrophoresis in 1.0% agarose gels and detected by staining with ethidium bromide.

For the PCR amplification preceding the pyrosequencing reaction, initial incubation was at 94°C for 2 min, followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 57°C for 30 sec, and polymerization at 72°C for 20 sec. Final polymerization at 72°C was carried out, for 5 min, after completion of the above cycles.

Mapping the DDB1 gene

DDB1 was mapped by means of *Lycopersicon (L.) pennellii* introgression lines (Eshed et al. 1992). DNA extracted from individual plants of each of the introgression lines, including their original parental lines M82 and *L. pennellii*, were used as templates in PCR reactions. The primers used for these mapping reactions were mTDDB F and mTDDB R (Table 1). These primers were derived from the Institute of Genomic Research (TIGR) database accession TC117372 (http://www.tigr.org/) that was found highly homologous to both copies of the *A. thaliana DDB1* gene. To obtain polymorphism between M82 and *L. pennellii*, the PCR products were digested with *Pst* I endonuclease, following the PCR reaction.

Cloning and sequencing of the tomato DDB1 cDNA from hp-1 and hp-1^{w} mutant plants

Total RNA was extracted from 25 mg of leaf tissue of individual *hp-1* and *hp-1*^{w} mutant seedlings and their nearly isogenic open pollinated wild-type genotypes (Ailsa Craig and GT, respectively). The RNA extraction was carried out using the TRIzol reagent system (GibcoBRL Life Technologies, Gaithersburg, MD, USA). Total RNA was used as the template for first-strand cDNA synthesis using the Superscript pre-amplification system (GibcoBRL Life Technologies, Paisley, UK). The cDNA prepared was used as a template in PCR reactions, to amplify overlapping fragments of the gene encoding the tomato *DDB1*, from in both mutant and normal genetic accessions. The PCR products were then sequenced, either directly or after cloning into pGEM-T Easy Vector using the pGEM-T Easy Vector Systems, according to the manufacturer recommendations (Promega, Madison, WI, USA). After cloning into pGEM-T Easy Vector, four or five independent clones of each of the overlapping amplified fragments were sequenced, based on the vector T7, SP6, and primers complementary to the tomato *DDB1* gene. Whenever direct sequencing was used, at least two PCR products, representing each primer combination complementary to the tomato *DDB1* gene, were sequenced. Sequencing was carried out with an ABI PRISM 377 automated DNA sequencer (Applied Biosystems, Foster City, CA, USA).

The 3' region of the tomato *DDB1* gene was directly sequenced by using overlapping fragments amplified with primers complementary to TIGR data base accession TC117372 (http://www.tigr.org/) that is highly homologous to both copies of the A. thaliana *DDB1* gene. These primers are presented in Table 1, below:

Table 1. Forward (F) and reverse (R) primers, complementary to TIGR database accession TC117372, used to sequence the 3' region of the tomato *DDB1* gene.

Primer name	Primer sequence	Sequence list ref. no.
5TDDB F	5'-ACGACCTATCGTGGACTTCTGT-3'	SEQ ID NO:3
5TDDB R	5'-CTGGACTTGAGAATTGAAGCCT-3'	SEQ ID NO:4
In5TDDB F	5'-GAGCCTATAAGGATGGATCAC-3'	SEQ ID NO:5
ATDDB F	5'-CAGCAGTTGGAATGTGGACAG-3'	SEQ ID NO:6
mTDDB F	5'-GCAATCGCTAAAGAAGGTGAGT-3'	SEQ ID NO:7
mTDDB R	5'-GCATTATAGTCTCTGGCTCGCT-3'	SEQ ID NO:8
inmTDDB F	5'-GGACATTTGCTCTATGCAGT-3'	SEQ ID NO:9
inmTDDB R	5'-AGGCATTTAGAGAGTAGACAGC-3'	SEQ ID NO:10
TDDB F	5'-TTTGGAGAAGCTGCAGACAA-3'	SEQ ID NO:11
TDDB R	5'-CACAACCTCACAGAAGAAGAAG-3'	SEQ ID NO:12
In3TDDB R	5'-CCACTCTCTTCATTAGTTCCTC-3'	SEQ ID NO:13

The 5' region of the *DDB1* gene was initially cloned from a pBluescript® SK(+/-) phagemid cDNA library with the following primers:

T7= 5'-GTAATACGACTCACTATAGGGC-3' (SEQ ID NO:14)and

5'TDDB_R= 5'-CTGGACTTGAGAATTGAAGCCT-3'

This cDNA library, kindly provided by R. Barg and Y. Salts, of the Volcani Center, Israel, was prepared from young parthenocarpic fruits of 4-6 mm in diameter (ca. 4-8 days post-anthesis) derived from the facultative parthenocarpic determinate line L-179 (*pat-2/pat-2*). This line was described previously (Barg et al. 1990). The library was prepared with the cDNA Synthesis Kit#200400, Zap-cDNA Synthesis Kit#200401, and Zap cDNA Gigapack III Gold Cloning Kit#200450 of Stratagen Inc., according to the of the manufacturer's instructions.

The 5' region of the tomato *DDB1* gene from *hp1/hp1* and *hp1^w/hp1^w* mutant lines and their corresponding nearly isogenic normal lines was directly sequenced using the above primer (5'TDDB_R) and the primer TDB_UTR= 5'-

ATAGCGGGAAGAGGGAAGATAC-3' (SEQ ID NO:15), that is complementary to the 5' UTR of the tomato *DDB1* gene. Several overlapping primers complementary to the above fragment, such as those used for pyrosequencing genotyping (see below), were used for sequence verification of the 5' coding sequence of the tomato *DDB1* gene.

Linkage analysis

The analysis of linkage between the tomato *DDB1* locus and the exaggerated photomorphogenic de-etiolation response characterizing *hp-1* mutant, was carried out using F_2 seeds of a cross between *hp-1* mutant plants and wild-type plants (cv. Ailsa Craig). These seeds were allowed to germinate under a yellow plastic screen that prevented the transmission of light of wavelengths under 500nm (Mochizuki and Kamimura 1984), in an environmentally controlled growth chamber (25°C day/18°C night). These germination and initial growth conditions result in exaggeration of hypocotyl-length differences between the mutant and normal plants (Mochizuki and Kamimura, 1984). The hypocotyl lengths of individual F_2 seedlings were measured 8 days after sowing, and their genotype was determined with the pyrosequencing-based DNA marker disclosed and described herein.

Pyrosequencing genotyping

A pyrosequencing genotyping system (extensively reviewed by Ronaghi 2001) based on the above-described single nucleotide polymorphism (SNP) between *hp1/hp1* mutant line and its nearly isogenic normal line in cv. Ailsa Craig background was developed. For this purpose a genomic fragment containing the SNP was cloned and sequenced as presented in Fig. 1. The biotin-labeled forward primer for this reaction was 5'-TGTTTTCCAGAGTTACCGGACT-3' (SEQ ID

NO:16); the reverse primer was 5'-TAGCTTGAGCCAATGAAGACAA-3' (SEQ ID NO:17); and the sequencing primer was 5'-ATGAAGACAAAAGCAT-3' (SEQ ID NO:18). The amplicon size in this reaction was 106 bp.

The PCR amplification reaction preceding the pyrosequencing reaction was as described above (see PCR reactions). Two pmoles of the sequencing primer were added to the amplification reaction prior to the pyrosequencing analysis. The analysis was carried out using a MegaBASE 1000 instrument (Danyel Biotech, Nes Ziona, Israel). Because the sequencing primer is in reverse orientation, the normal genotype is characterized by T whereas the homozygous mutant *hp-1* genotype is characterized by A at the SNP location, as shown in Fig. 3.

Statistical analyses

Analyses of variance (ANOVA) were carried out with the JMP Statistical Discovery software (SAS Institute, Cary, NC, USA). Linkage analysis and LOD score determination were carried out with the QGENE software Version 3.06d (Nelson 1997). Alignment of amino-acid sequences was carried out using the Clustal method (Higgins and Sharp 1988).

Example 1

Identification and cloning of the tomato homolog of *DDB1*

The DDB1 protein is a heterodimer consisting of two subunits, DDB1 and DDB2. Unlike rice, chicken, human, mouse, *Drosophila* and *Schizosaccharomyces pombe*, the *A. thaliana* genome harbors two highly homologous copies of the *DDB1* gene (Schroeder et al. 2002; Zolezzi et al. 2002; Fu et al. 2003; Ishibashi et al. 2003): DDB1A, and DDB1B, both 1088 amino-acids in length (Genbank protein accessions

NP_192451 and NP_193842, respectively). When each of these two protein accessions were used as a query in tblastn analysis against the TIGR database (http://www.tigr.org/) containing tomato Expressed Sequence Tags (EST), both revealed two highly homologous sequences: TC117371 (394 bp) and TC117372 (2206 bp). The *A. thaliana* Accession NP_192451 was found to share 87 and 86% identities with the tomato TC117371 and TC117372 accessions, respectively. Accession NP_193842, on the other hand, shared 87 and 83% identities with the tomato TC117371 and TC117372 accessions, respectively. Careful sequence analysis, based initially on the longer TIGR accession, TC117372, and later on the single gene that we had cloned from a cDNA library, made it clear to us that the two tomato TIGR accessions, TC117371 and TC117372, were complementary to the same gene sequence. Moreover, Southern-blot transfer and hybridization of tomato genomic DNA, with the *DDB1* gene sequence as a probe, revealed that indeed the tomato genome contains a single copy of the *DDB1* gene (data not presented).

Example 2

Mapping of the tomato *DDB1*

Partial mapping results, that include the approximate map location of the tomato *DDB1* gene, are presented in Fig. 2. These results indicate that the *DDB1* is located on the tomato chromosome 2, in the introgression line that harbors the *HP-1* gene (Yen et al. 1997).

Example 3

Sequence characterization of the tomato *DDB1* in *hp-1* and *hp-1^{w}* mutants

Several forward and reverse primers (Table 1), complementary to the 3' region of the tomato DDB1 gene (TIGR accession TC117372), were used in order to perform direct sequencing on cDNA prepared from leaves of seedlings from *hp-1* and normal plants in Ailsa Craig background. No polymorphism was obtained between *hp-1* and normal plants in this region. The 5' region of the DDB1 gene in the two genotypes was therefore cloned and thoroughly sequenced as well. Computerized translation of all sequence results showed that the tomato DDB1 is a 1090-amino-acid protein. Sequence analysis of the *DDB1* coding sequence from *hp-1* and its nearly isogenic normal genotype revealed a single A^{931}-to-T^{931} base transversion in the coding sequence of *DDB1* gene of the mutant *hp-1* plants. This transversion resulted in a substitution of a conserved $Asparagine^{311}$ to $Tyrosine^{311}$ (Fig. 4).

Based on the sequence information obtained in the Ailsa Craig background, we have also sequenced the entire coding region of the *DDB1* gene in *hp-1^{w}* mutant plant and its isogenic normal counterpart in GT background. Because *hp-1^{w}* is allelic to *hp-1*, a major mutation in the coding sequence of the *DDB1* gene in *hp-1^{w}* mutants would strongly support the hypothesis that the tomato *DDB1* gene causes both the *hp-1* and *hp-1^{w}* mutant phenotypes. Indeed, a single G^{2392}-to-A^{2392} transition was observed in the *DDB1* coding sequence in the *hp-1^{w}* mutant plant which results in a substitution of a conserved $Glutamic\text{-}acid^{798}$ to $Lysine^{798}$ (Fig. 4).

The complete nucleotide coding sequence and the deduced amino acids sequence of the normal wild-type tomato *DDB1* gene are shown in Fig. 5 and Fig. 6, respectively.

Example 4

Genotyping of lines and cultivars

Nineteen lines or cultivars, obtained from various sources, were genotyped by a combination of direct sequencing and pyrosequencing methods (Fig. 3). Included among them were a single heterozygous *hp-1/+*, 10 *hp-1/hp-1* and eight normal +/+ accessions. Complete agreement between the SNP identified at the *DDB1* gene and the known genotype of the plants at the *HP-1* locus was found (results not shown).

Example 5

Analysis of the linkage between the *DDB1* locus and the photomorphogenic response

A linkage analysis study was carried out to test the association between the *DDB1* locus and the characteristic hypersensitive-photomorphogenic response displayed by *hp-1* mutant seedlings (i.e., the inhibition of hypocotyl elongation phenotype). For this purpose, F_2 seeds of a cross between determinate *hp-1* mutant plants and wild type plants (cv. Ailsa Craig) were germinated under yellow plastic screen in a controlled growth chamber. Eight days after sowing, the hypocotyls lengths of individual seedlings were recorded, and their *DDB1* locus was genotyped with aid of the pyrosequencing DNA marker, as described above. The results demonstrate a clear association between

the *DDB1* locus and hypocotyls length, as shown in Table 2, below:

Table 2. Linkage analysis between the tomato *DDB1* locus and the photomorphogenic response displayed by *hp-1* mutant seedlings. Seedlings were grown under a yellow plastic screen for 8 days after sowing. Different superscript letters indicate statistically significant differences between means ($P<0.05$) according to the Tukey-Kramer HSD test (Kramer 1956).

Genotype	N	Hypocotyl length ± S.E. (cm)	LOD score	R^2
+/+	35	9.6^A ±0.2	25<LOD<26	62.8%
hp-1/+	68	8.7^B ±0.2		
hp-1/hp-1	20	4.2^C ±0.2		

Homozygous recessive *hp-1/hp-1* seedlings displayed a highly significant inhibition of hypocotyl elongation, indicative of a more exaggerated photomorphogenic de-etiolation response, in comparison to the two other two genotypic groups (25<LOD Score<26, R^2= 62.8%). These results confirm that the mutation identified in the *DDB1* locus of *hp-1* mutant plants is associated with one of its main characteristic phenotypes, i.e., inhibited hypocotyl elongation in the seedlings. Interestingly, a slight partially dominant effect for the *hp-1* allele was obtained in this study. This effect can be noted from the statistically significant differences obtained between the +/+ and *hp-1*/+ group means (Table 2).

Further non-limiting examples (both working and theoretical) that illustrate and describe various practical embodiments of the present invention are given in the following sections. These embodiments are described for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

Example 6

Diagnostic tool for identifying the *hp-1 and hp-1^{w}* mutations

A pyrosequencing DNA marker system, extensively reviewed by Ronaghi 2001, for use as a molecular diagnostic tool for identifying *hp-1* mutant plants based on the sequence results (Fig. 1) was developed. This DNA marker is based on the single nucleotide polymorphism (SNP) discovered in this study between *hp1/hp1* mutant line and its nearly isogenic normal line in cv. Ailsa Craig background. For this purpose a genomic fragment containing the SNP was cloned and sequenced. The sequence of this genomic fragment is presented in Fig. 1. The biotin labeled forward primer for this reaction was: 5'-TGTTTTCCAGAGTTACCGGACT-3' (SEQ ID NO:16); the reverse primer was: 5'-TAGCTTGAGCCAATGAAGACAA-3' (SEQ ID NO:17); and the sequencing primer was 5'-ATGAAGACAAAAGCAT-3' (SEQ ID NO:18). The amplicon size in this reaction was 106 bp.

The PCR amplification reaction preceding the pyrosequencing reaction was as described (see PCR reactions above). Two pmoles of the sequencing primer were added to the amplification reaction prior to the pyrosequencing analysis. The analysis was carried out using a MegaBASE 1000 instrument by Danyel Biotech, Nes Ziona, Israel. Because the sequencing primer is in reverse orientation, the normal genotype is characterized by T whereas the

homozygous mutant *hp-1* genotype is characterized by A at the SNP location as presented in fig. 3.

Using the pyrosequencing methodology and the primers described above a clear polymorphism between *hp-1* and wild-type plants was seen as demonstrated in Fig. 3. In the case of the homozygous *hp-1* mutant plants, a single peak representing Adenine (A) at the SNP location was observed, while in wild-type plants, a single peak representing Thymine (T) at the SNP location was observed (Fig. 3). As expected, plants heterozygous for the *hp-1* mutation yielded two peaks, representing both A an T nucleotides (Fig. 3).

A similar pyrosequencing based marker system based on the SNP observed in *hp-1*^{w} mutant plants has also been established.

Example 7

Incorporation of two genetically unlinked lycopene enhancing mutations in a single tomato hybrid: *Experimental Approach*

A common practice among breeders is to combine or incorporate two or more mutations positively affecting the same trait. Such procedure can be verified by laborious and time consuming test crosses. The diagnostic tool produced herein can facilitate the incorporation of two light hypersensitive lycopene-enhancing mutations in a single plant or breeding line.

Several mutations in tomato positively affect lycopene content in the mature tomato fruit. Of these, at least 5

show a significant hypersensitive light response. These include:

1. *High pigment-1* (*hp-1*)
2. *High pigment-1^{w}* (*hp-1^{w}*)
3. *High pigment-2* (*hp-2*)
4. *High pigment-2^{j}* (*hp-2^{j}*)
5. *Dark green* (*dg*)

The *hp-1* and *hp-1^{w}* mutations map to the *HP-1* locus on the tomato chromosome 2 (Yen et al. 1997 and in accordance with the present invention). The *hp-2*, *hp-2^{j}* and *dg* mutations map to the *HP-2* locus on the tomato chromosome 1 (Mustilli et al. 1999; Levin et al. 2003). Incorporation of lycopene enhancing *hp-2*, *hp-2^{j}* or *dg* at the *HP-2* locus and either one of the two mutations that map to the *HP-1* locus (*hp-1* and *hp-1^{w}*) can be more efficiently achieved through the following procedure (illustrated for the *dg* and *hp-1* mutations):

1. Cross homozygous *dg* with homozygous *hp-1* mutants to generate double heterozygous F_1 plants:

dg/dg +/+ X +/+ *hp-1/hp-1*

↓

dg/+ hp-1/+

2. Self-cross the F_1 double heterozygous plants to generate F_2 seeds. These F_2 seeds will segregate into 9

genotypes: *dg/dg hp-1/hp-1, dg/dg hp-1/+, dg/dg +/+, dg/+ hp-1/hp-1, dg/+ hp-1/+, dg/+ +/+, +/+ hp-1/hp-1, +/+ hp-1/+, +/+ +/+.*

Using the pyrosequencing marker system for the *hp-1* mutation disclosed herein and the marker for the *dg* mutation disclosed in co-owned pending application PCT/IL03/00023, the double homozygous plants *dg/dg hp-1/hp-1* can be easily identified and self-crossed to yield a breeding line homozygous for the two mutations.

Example 8

Incorporation of two genetically unlinked lycopene enhancing mutations in a single tomato hybrid significantly increases lycopene yield: *Working Example*

Two semi-isogenic hybrids, one homozygous for the *hp-1* mutation, *hp-1/hp-1*, and the other for the *dg* mutation, *dg/dg*, were crossed hybridized to yield F_1 plants (*hp-1/+ dg/+*). These F_1 plants were self-hybridized to yield F_2 seedlings. These F_2 seedlings were genotyped and self-hybridized to yield double mutant plants (*hp-1/hp-1 dg/dg*), as outlined in Example 7. Two horticulturally acceptable plants were selected and allowed to self hybridize to yield two F_4 lines. These F_4 lines were cross hybridized to yield a double mutant hybrid. This hybrid was tested, together with the semi-isogenic single mutant hybrids used in the initial cross (see above), in 4 locations in northern Israel during the spring season under open field conditions. Results presented in Table 3 (below) show that, unexpectedly, the lycopene yield of the double mutant hybrid is statistically higher compared to its isogenic single mutant hybrids. The increase in lycopene yield of

the double mutant hybrid was 19 and 61% compared to the lycopene yield of the *dg/dg* and *hp-1/hp-1* single mutant hybrids, respectively.

Table 3. Lycopene yield of single and double mutant hybrid cultivars carrying hypersensitive lycopene-enhancing mutations. Different superscript letters represent statistically significant differences between means ($P<0.05$) based on Tukey-Kramer HSD test (Kramer, 1956).

Cultivar genotype	Lycopene yield (gr/dunam*)
+/+ hp-1/hp-1	1136^{C}
dg/dg +/+	1538^{B}
dg/dg hp-1/hp-1	1824^{A}

* 1 dunam= 1000 square meters

Example 9

Use of the diagnostic tool for post control analysis of parental lines and hybrid seeds

Seed companies often use a battery of molecular markers for post- or quality- control of parental seed stocks and hybrid-seeds. Several commercial lycopene-rich tomato cultivars carry the *hp-1* and *hp-1^{w}* mutation either at a homozygous or heterozygous state. Up until now, detection of the *hp-1* and *hp-1^{w}* traits within a particular stock could only be performed by the lengthy procedure of germinating samples of the seeds, and performing complicated phenotypic analyses on the parental cultivars and subsequent generations.

The diagnostic tool demonstrated in this study (see Example 6, hereinabove) can be used to positively detect the *hp-1* and *hp-1*^{w} alleles in such cultivars and their parental lines, and thus enable post-production quality control to be carried out over a time scale of 1-2 days instead of weeks or months.

Example 10

Mapping of other functionally active mutations in the *DDB1* gene

Seeds extracted from normal plants can be mutagenized with ethyl methanesulfonate (EMS) or other approaches according to known protocols to yield photomorphogenic mutants (Koornneef et al. 1990). These mutants can be selected for under modulated light conditions, such as yellow plastic screen. Photomorphogenic mutants obtained can be screened for unique expression patterns of health-promoting metabolites. These mutants can be further characterized by allele tests against *hp-1* and/or *hp-1*^{w} and some of them may be characterized as allelic to these mutations. Thus, mutagenized plants allelic to *hp-1* and/or *hp-1*^{w} can be discovered that also bear unique metabolite profiles. Sequence analysis of the *DDB1* gene in these plants should reveal the exact genetic modifications that underline such unique metabolic architectures. These genetic modifications should enable the design of specific molecular markers, similar to those outlined hereinabove, for marker assisted selection. Also, mapping of such lesions may unravel regions within the *DDB1* gene as targets for efficient genetic manipulation to obtain plants with unique metabolite profiles in the tomato fruit.

Example 11

Over-expression of normal or modified *DDB1* genes to obtain over-production of health promoting metabolites in the tomato fruit and/or fruits and vegetables of other plant species

The DDB1 gene is highly conserved across many evolutionary distant species (Schroeder et al. 2002). Also, its link to overproduction of heath-promoting metabolites has been outlined hereinabove. These results suggest that effects of the *DDB1* gene on the production of health promoting compounds should not be ignored in other plant species as well. From such practical point of view, *the DDB1* gene may be cloned from normal or *hp-1* and *hp-1*^{w} mutant tomato plants, or any other plant species, in sense or anti-sense (RNAi) orientations under constitutive or fruit-specific promoters. Over-expression of any of these constructs in any plant species may result in increases in the production of functional metabolites in fruits and vegetables.

References

Barg R, Meir E, Lapushner D, Frankel R, Salts Y (1990) Differential regulation of fruit specific 62 kDa protein in developing parthenocarpic (pat-2/pat-2) and seeded tomato fruits. *Physiol Plant* 80:417-424

Chory J (1993) Out of darkness: mutants reveal pathways controlling light-regulated development in plants. Trends Genet 9:167-172

Eshed Y, Abu-Abied M, Saranga Y, Zamir D (1992) Lycopersicon esculentum lines containing small overlapping introgressions from L. pennellii. Theor Appl Genet 83:1027-1034

Fu D, Wakasugi M, Ishigaki Y, Nikaido O, Matsunaga T (2003) cDNA cloning of the chicken DDB1 gene encoding the p127 subunit of damaged DNA-binding protein. Genes Genet Syst 2003 78:169-77

Fulton TM, Chunwongse J, Tansley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13:207-209

Higgins DG, Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237-244.

Ishibashi T, Kimura S, Yamamoto T, Furukawa T, Takata K, Uchiyama Y, Hashimoto J, Sakaguchi K (2003) Rice UV-damaged DNA binding protein homologues are most abundant in proliferating tissues. Gene 308:79-87

Konsler TR (1973) Three mutants appearing in 'Manapal' tomato. HortSci 8:331-333

Koornneef M, Bosma TDG, Hanhart CJ, van der Veen JH, Zeevaart JAD (1990) The isolation and characterization of gibberellin-deficient mutant in tomato. Theor Appl Genet 80:852-857

Kramer CY (1956) Extension of multiple range tests to group means with unequal number of replications. Biometrics 12:309-310

Levin I, Frankel P, Gilboa N, Tanny S, Lalazar A (2003) The tomato dark green mutation is a novel allele of the tomato homolog of the DEETIOLATED1 gene. Theor Appl Genet 106:454-460

Levin I, Smith EJ (1990) Molecular analysis of endogenous virus ev21-slow feathering complex of chickens. 1. Cloning of proviral-cell junction fragment and unoccupied integration site. Poult Sci 69:2017-2026

Mochizuki T, Kamimura S (1984) Inheritance of vitamin C content and its relation to other characters in crosses between *hp* and *og* varieties of tomatoes. Eucarpia Tomato Working Group. Synopsis IX meeting 22-24 May, Wageningen, the Netherlands, pp 8-13

Mustilli AC, Fenzi F, Ciliento R, Alfano F Bowler C (1999) Phenotype of the tomato *high pigment-2* is caused by a mutation in the tomato homolog of *DEETIOLATED1*. Plant Cell 11:145-157

Nelson JC (1997) QGENE: software for marker-based genomic analysis and breeding. Mol Breed 3:229.235

Peters JL, van Tuinen A, Adamse P, Kendrick RE, Koornneef M (1989) High pigment mutants of tomato exhibit high sensitivity for phytochrome action. J Plant Physiol 134:661-666

Reynard GB (1956) Origin of Webb Special (Black Queen) in tomato. Rep Tomato Genet Coop 40: 44-64

Ronaghi M (2001) Pyrosequencing sheds light on DNA sequencing. Genome Res 11:3-11

Schroeder DF, Gahrtz M, Maxwell BB, Cook RK, Kan JM, Alonso JM, Ecker JR, Chory J (2002) De-etiolated 1 and damaged DNA binding protein 1 interact to regulate Arabidopsis photomorphogenesis. Curr Biol. 12:1462-1472

Soressi GP (1975) New spontaneous or chemically-induced fruit ripening tomato mutants. Rep. Tomato Genet Coop 25:21-22

van Tuinen A, Cordonnier-Prat M-M, Pratt LH, Verkerk R, Zabel P, Koornneef M (1997). The mapping of phytochrome genes and photomorphogenic mutants of tomato. Theor Appl Genet 94:115-122

Wann EV, Jourdain EL, Pressey R Lyon BG (1985) Effect of mutant genotypes *hp* og^c and *dg* og^c on tomato fruit quality. J Amer Soc Hort Sci 110:212-215

Yen H, Shelton A, Howard L, Vrebalov J, Giovanonni JJ (1997) The tomato *high-pigment* (*hp*) locus maps to chromosome 2 and influences plastome copy number and fruit quality. Theor Appl Genet 95:1069-1079

Zolezzi F, Fuss J, Uzawa S, Linn S (2002) Characterization of a Schizosaccharomyces pombe strain deleted for a sequence homologue of the human damaged DNA binding 1 (DDB1) gene. J Biol Chem 277:41183-411891